WPI

Wireless Innovation Laboratory

Multi-Objective Reinforcement Learning for Cognitive Radio-based Satellite Communications

Paulo Ferreira, Randy Paffenroth, Alexander Wyglinski - *(WPI)*

Timothy Hackett, Sven Bilen - *(Penn State)*

Richard Reinhart, Dale Mortensen - *(NASA GRC)*

NASA GRC Grant: "Intelligent Media Access Protocol for SDR-based Satellite Communications."

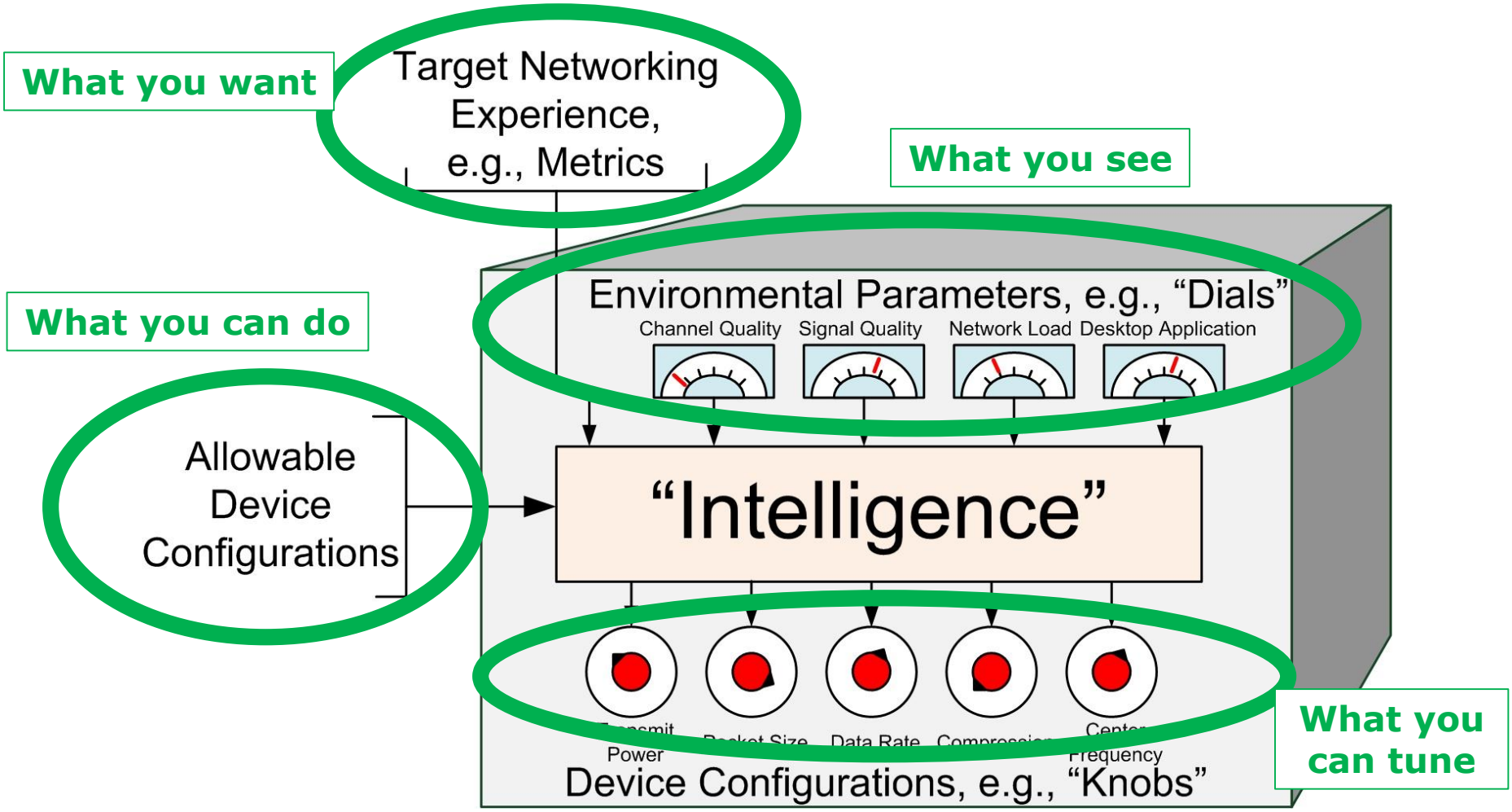
Grant number: NNC14AA01A

Agenda

- What is a Cognitive Radio?
- CR applications
- The problem: Multi-objective performance
- Reinforcement Learning: The solution
- Satcom RL performance

What is a Cognitive Radio ?

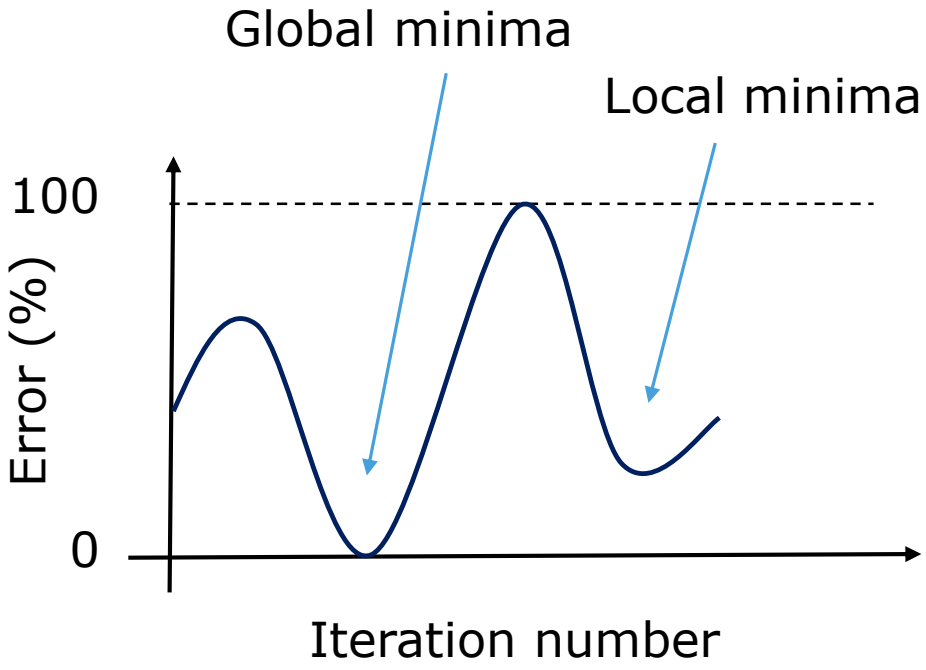
What is a Cognitive Radio ?



T. Collins, A.M. Wyglinski. *"Enabling Security in Cognitive Radios and Wireless Spectrum."* MILCOM Tutorial, 2014.

What is a Cognitive Radio ?

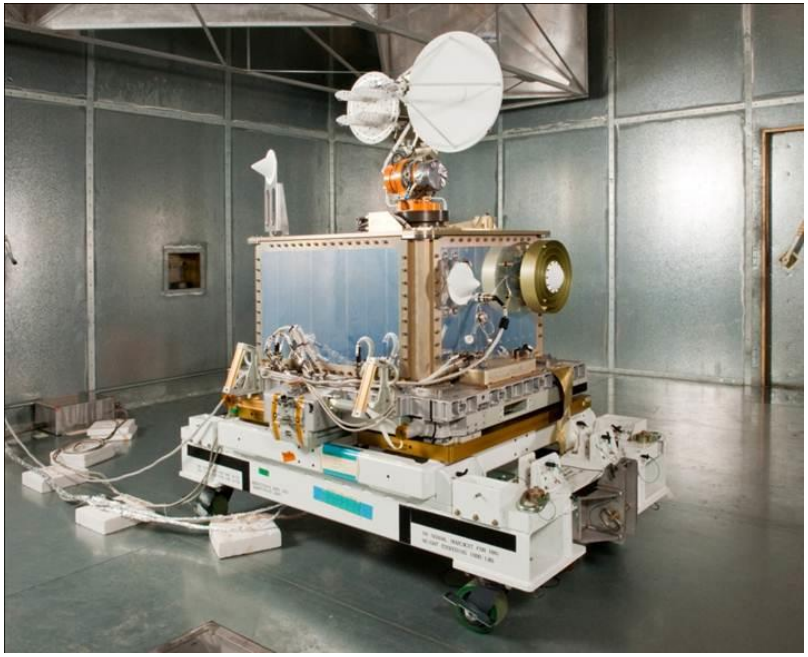
- Learning algorithm
  - Explore vs. Exploit



CR applications

CR applications

- **Satcom**



Reinhart, R. C. Using International Space Station For Cognitive System Research And Technology With Space-based Reconfigurable Software Defined Radios. 66th International Astronautical Congress, IAC 2015.

CR Satcom – What you tune: PHY layer

- Modulation scheme
- Encoding scheme
- Symbol rate
- Bandwidth
- Carrier frequency
- ADC/DAC resolution
- Antenna
- Transmission power level

Multi-Objective Comms Performance

The problem

Multi-objective comms performance

Table 1. Radio parameter adaptation inter-relationship

Adaptation	Goals	Consequences	Conflicts	Constants
$\downarrow M$	$\min(\text{BER})$, $\min(P)$	$\downarrow R$,P,BER	$\max(R)$	E_b
$\uparrow M$	$\max(R)$	$\uparrow R$,P,BER	$\min(\text{BER})$, $\min(P)$	E_b
$\uparrow R$	$\max(R)$, W const.*	$\uparrow W$,P	$\min(P)$, W const.*	M, E_b
$\downarrow R$	$\min(P)$, W const.*	$\downarrow W$,P	$\max(R)$, W const.*	M, E_b

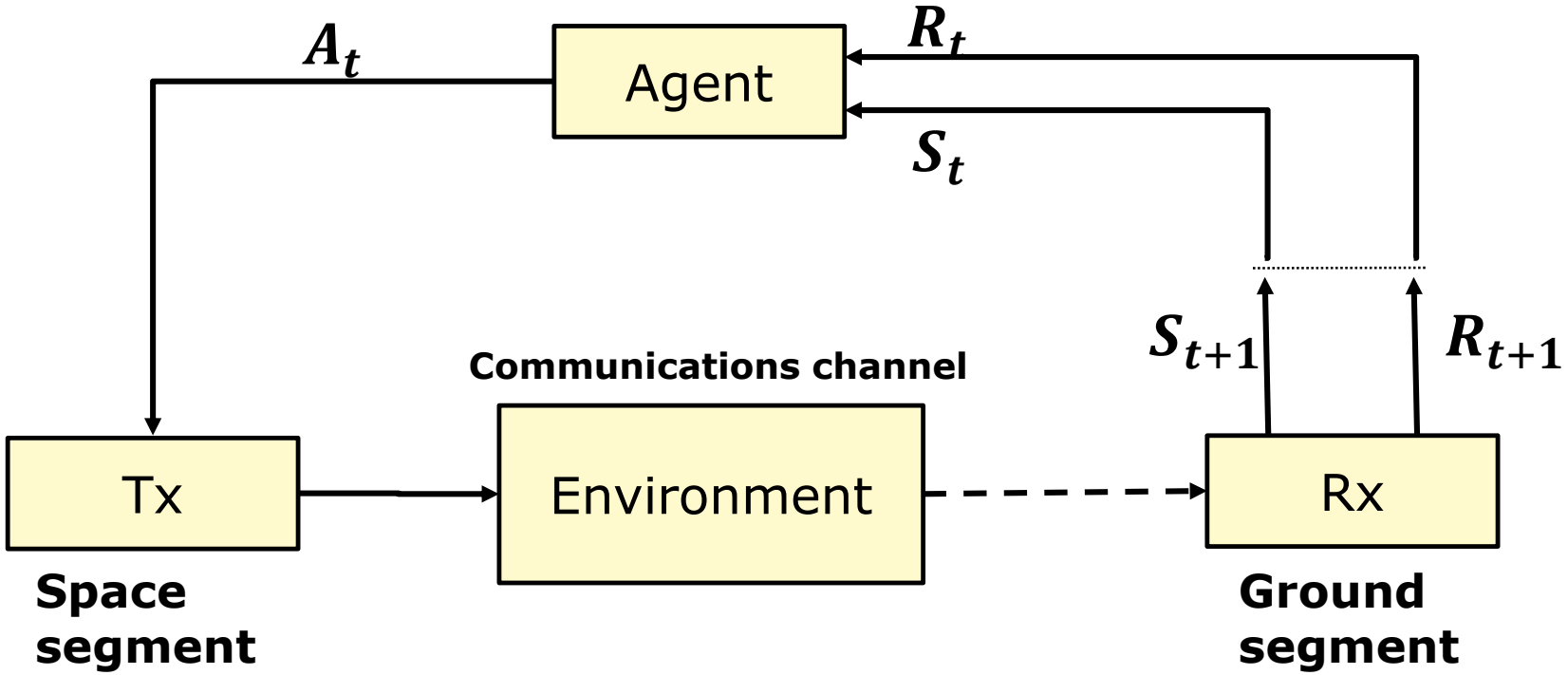
* Keeping W constant can be a goal or a conflicting goal while adapting R.

M – modulation and encoding schemes
R – data rate
BER – bit error rate
P – transmission power
W – bandwidth
Eb – energy per bit

Reinforcement Learning

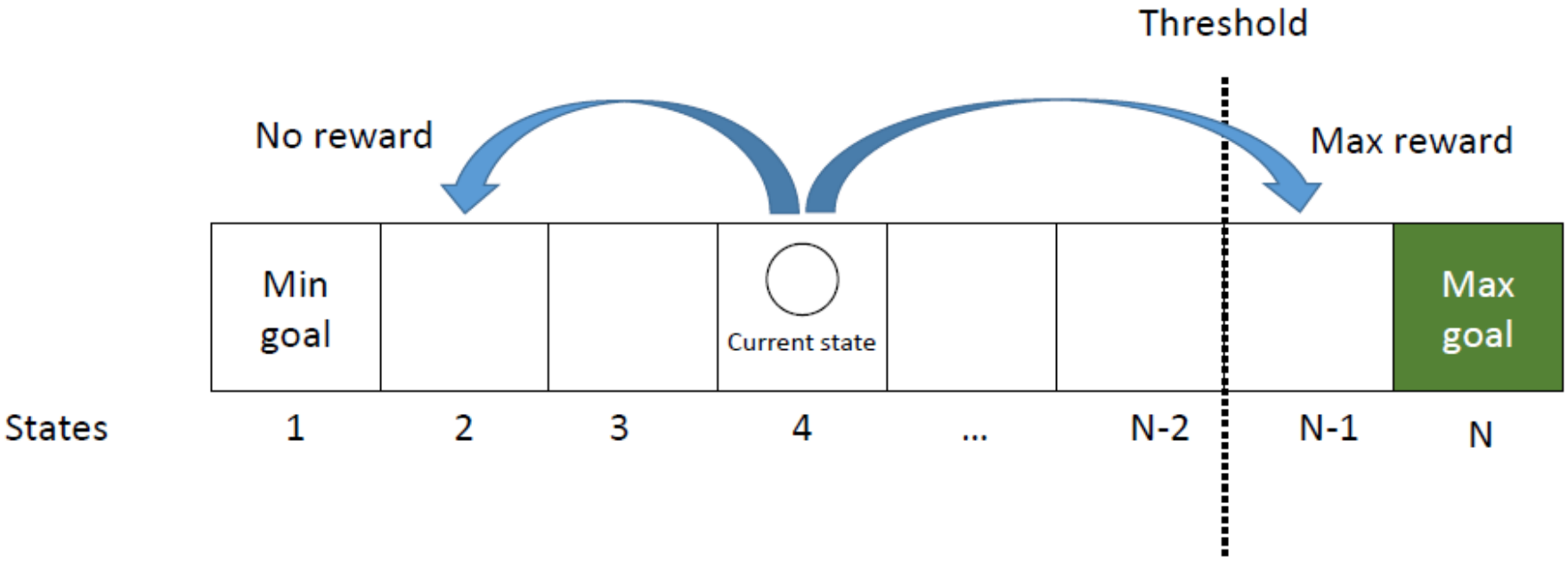
The solution

Reinforcement Learning



A_t=action
R_t= reward
S_t=state

Reinforcement Learning



Reinforcement Learning

- Satcom $\boldsymbol{Q(S,A)}$

$$Q_{k+1}(s_k, u_k) = Q_k(s_k, u_k) + \alpha_k r_{k+1}$$

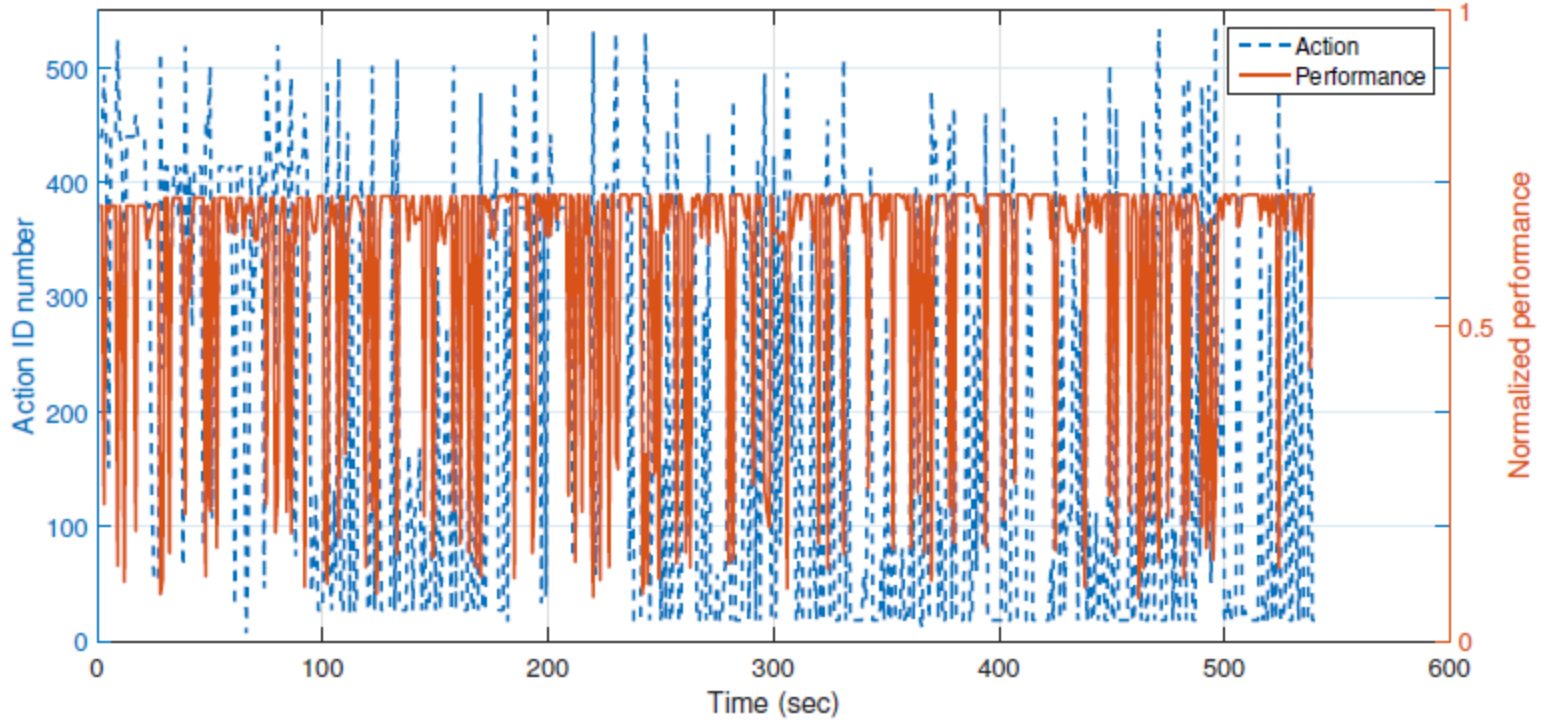
$$u_k = h(s_k) \longrightarrow \text{State-action policy}$$

$$s_{k+1} = g(s_k, u_k) \longrightarrow \text{State-transition function}$$

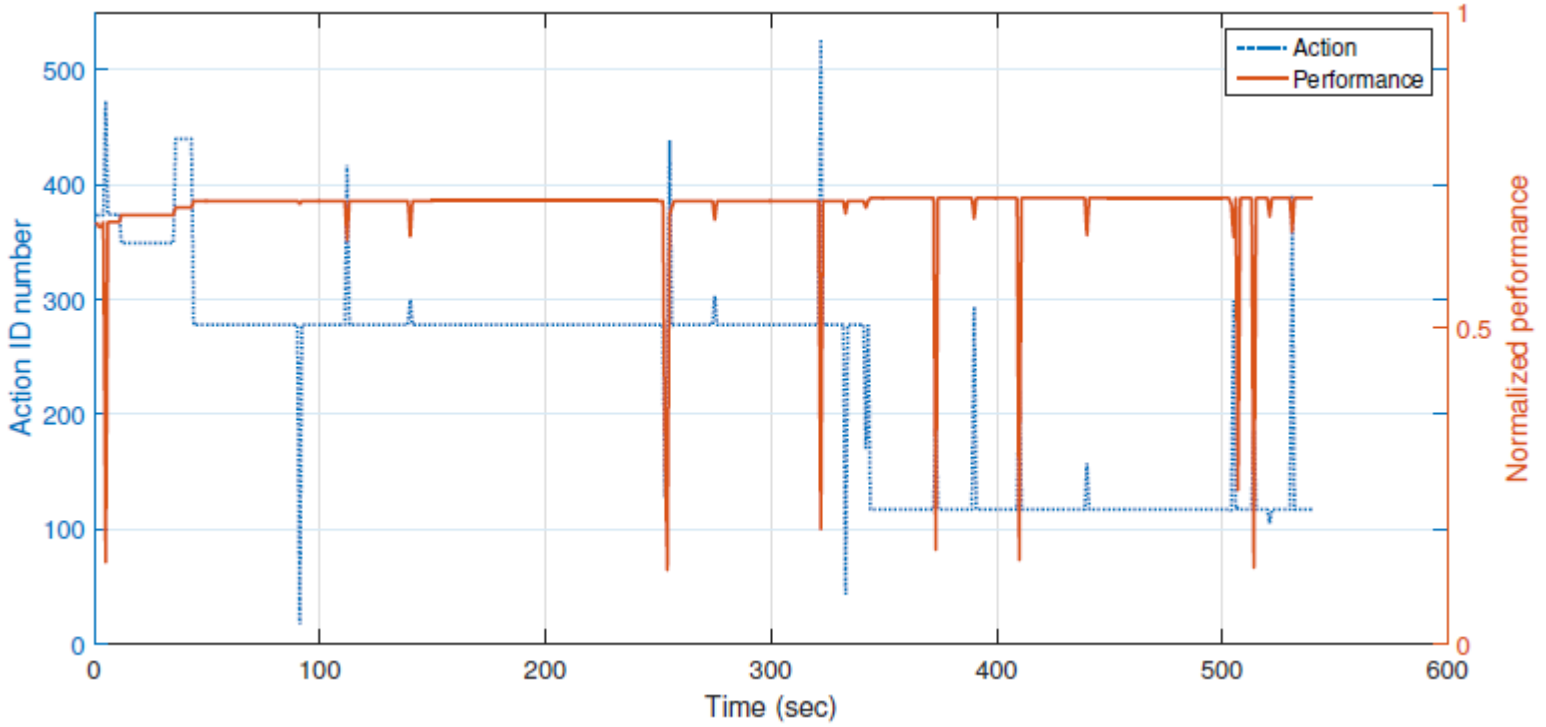
$$r_{k+1} = \rho(s_k, u_k) \longrightarrow \text{Reward function}$$

Satcom RL performance

Fixed exploration probability ($\varepsilon = 0.5$)

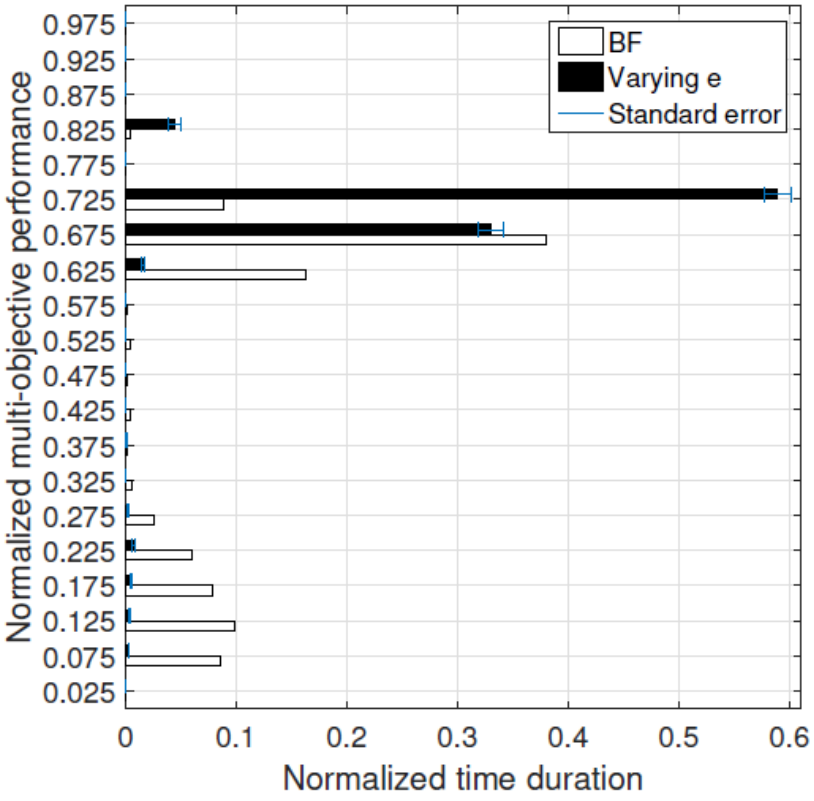


Variable exploration probability ε



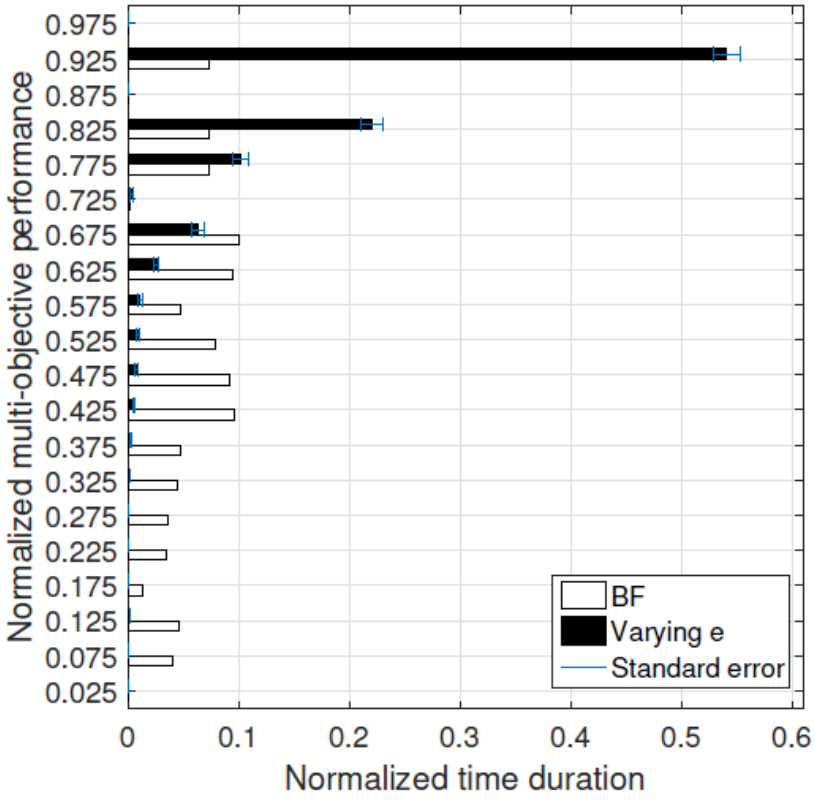
Time spent at performance levels

Mission 1: Launch/re-entry



(a)

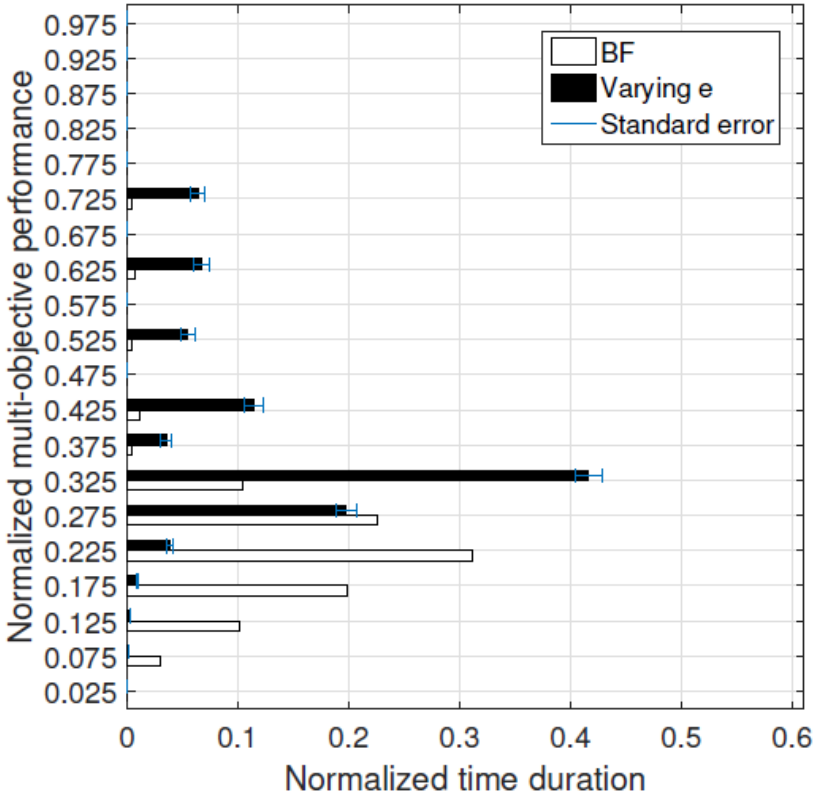
Mission 2: Multimedia



(b)

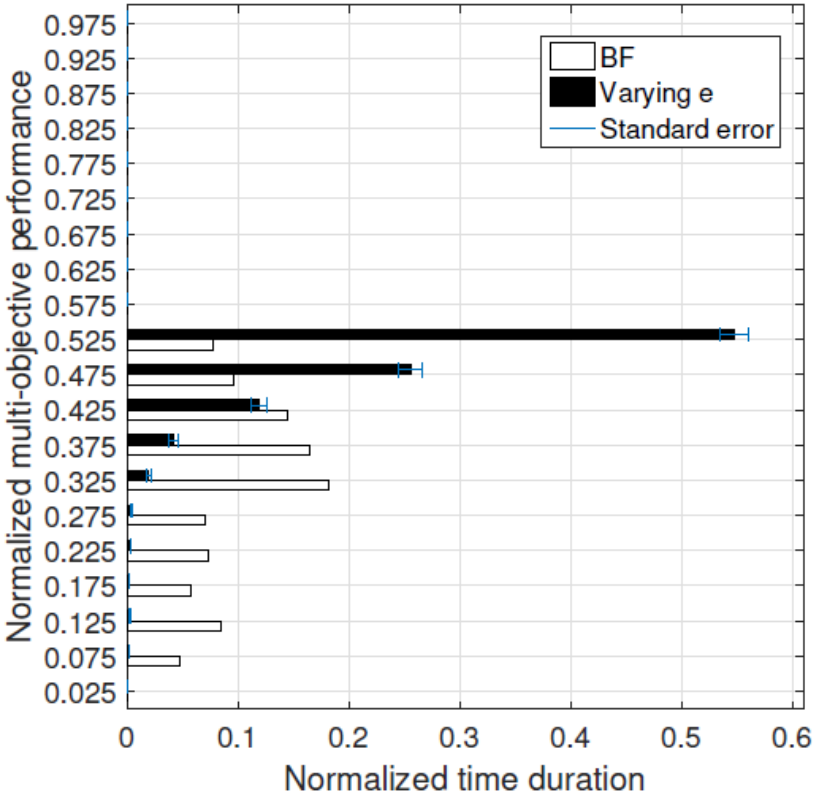
Time spent at performance levels

Mission 3: Power saving



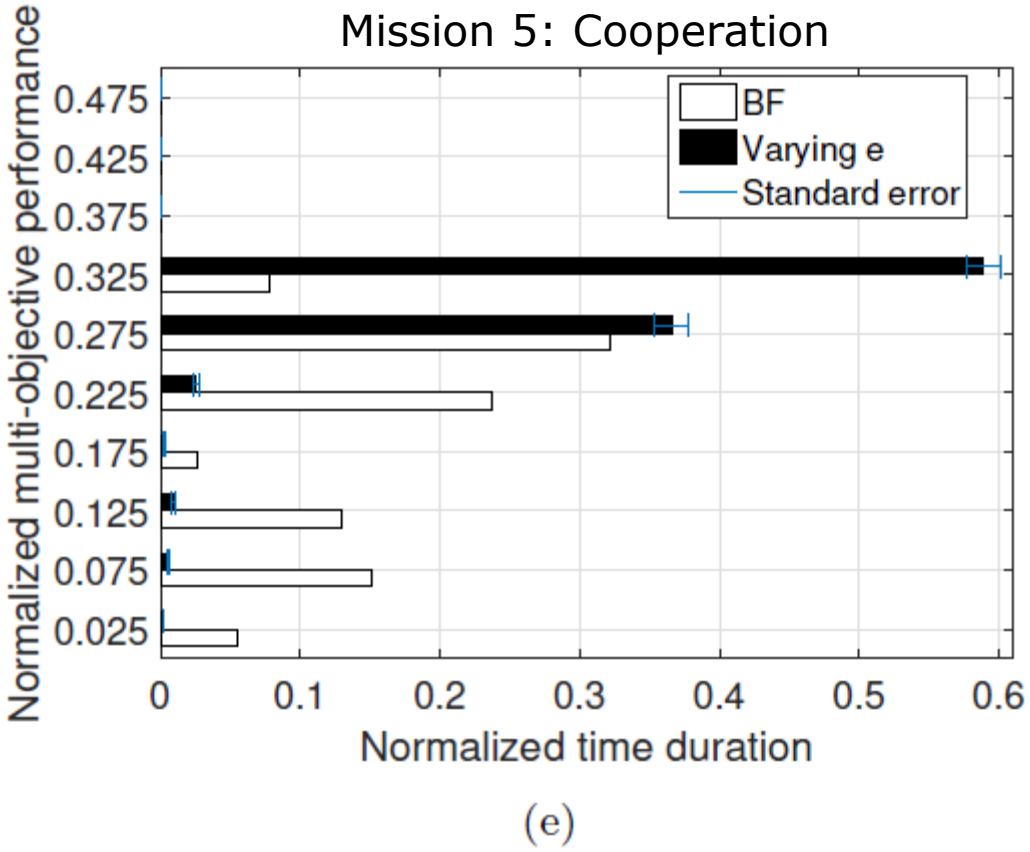
(c)

Mission 4: Normal



(d)

Time spent at performance levels



(e)

THANK YOU!

Wireless Innovation Laboratory

Applied Research Solutions for a Connected Society

Contact us

PRFERREIRA@wpi.edu

ALEXW@wpi.edu

http://www.wireless.wpi.edu/